Conceptual Change in Evolutionary Theory: The Effects of Scripted Argumentative Monologue in Peer Settings

Christa S.C. Asterhan (kooij@msschwarz.msc.huji.ac.il)
Baruch B. Schwarz (msschwarz@msschwarz.msc.huji.ac.il)
School of Education, Hebrew University of Jerusalem, Mt. Scopus
Jerusalem, 91905 Israel

Abstract

This study investigated the potentially beneficial role of argumentation on conceptual change in learning tasks that are based on the socio-cognitive conflict paradigm. Forty-two undergraduates participated in one of two conditions: Experimental students engaged in monological argumentation on their own and an equal-status peer confederate's explanation of an evolutionary phenomenon in response to prompts read by the confederate, whereas in the control condition subjects and the confederate merely read aloud their respective explanations. Students in the argumentative conditions were found to outperform control students on two different measures of conceptual understanding. The advantage of the argumentative conditions also held for the stricter criterion of conceptual change. The particular patterns of change are discussed. In addition, many students were found to be inconsistent in the explanatory schemas they applied at a given test occasion. We discuss the implications of this finding for the conceptualization and assessment of conceptual change.

Humans primarily construct personal conceptions about scientific phenomena on the basis of their personal experience in the world. These naive ideas are often incompatible with the scientific explanations pupils are confronted with in classroom settings. They may be replaced or restructured hopefully, but by no way inexorably, leading to the acquisition of correct scientific concepts. Such a process has traditionally been referred to as conceptual change (Vosniadou, 1999).

Research on inducing conceptual change among students has traditionally been conducted in the Piaget-inspired cognitive conflict paradigm in which students' naive conceptualizations are confronted with anomalous data or contradicting views or are paired with peers who have different views (socio-cognitive conflict). Inducing conceptual change has been extensively shown to be a hard goal to achieve, especially in individual settings (Limon, 2001): Students have to become aware of their initial (mis)conceptions, understand the new information presented to them, become aware of the contradictions, compare and evaluate the different ideas, and adapt their misconceptions.

In this paper, we present results from a study on the effects of engagement in dialectical argumentative reasoning on conceptual understanding in evolutionary theory. Specifically, we propose that the engagement in dialectical argumentation on one’s own and another person’s views by producing reasoned arguments in favour and against these ideas will facilitate conceptual change processes. We base this claim on the fact that dialectical argumentation intertwines a number of social and cognitive processes that are considered to promote concept learning:

First of all, similar to the 'self-explanation effect', the epistemic examination of one’s personal theories and the reasons behind them, has been known to promote understanding and knowledge construction processes (Chi, deLeeuw, Chiu & Lavancher, 1994; Kuhn, 1991). Furthermore, in dialectical argumentation participants are exposed to a multiplicity of ideas and encouraged to explore each other's validity. This implies that they have to consider objections to their personal theories and arguments, to attempt to understand alternative positions and to formulate objections and/or counter-objections. The unique structure of argumentation that links premises, conclusions, conditions, rebuttals and so forth is, in addition, thought to considerably improve and extend the organization of knowledge, leading to better recall and understanding on subsequent test occasions (Means & Voss, 1996).

In theory, the engagement in solitary argumentative reasoning may be expected to yield similar results. However, argumentation is basically a social process that presupposes the presence of an audience (Leitao, 2000). Solitary argumentation on scientific concepts is cognitively very demanding. The presence of a dialog partner may promote reflection and awareness to one's own beliefs (Amigues, 1988), cause learners to engage in explanatory activities (Okada & Simon, 1997) and reduce cognitive load through the personification of different alternatives.

In a recent study we investigated the effects of argumentative dialog on concept learning in evolutionary theory (Asterhan & Schwarz, 2005). Undergraduates were assigned to dyads and collaboratively tried to solve an evolutionary phenomenon (i.e., the evolution of webbed feet of ducks). Half of these dyads were instructed to engage in dialectical argumentative dialog on their respective solutions and received examples of argumentative moves within a dialog; the other half was merely instructed to collaborate. Individual evolutionary understanding was assessed on three test occasions: Prior to, immediately after and a week following the intervention. Delayed posttest explanations in the argumentative condition were found to be of a higher quality than in the control condition, when controlled for pretest performance. Furthermore, the pattern through of change revealed that students in both conditions improved their conceptual understanding immediately.
following the intervention; however, while students who were merely instructed to collaborate lost this temporary gain, students in the argumentative condition retained the same level of performance. Whereas manipulation check analyses were successful, not all experimental dyads engaged in dialectical argumentation, whereas a small number of control dyads spontaneously did. Post-hoc analyses on natural differences revealed that only dyads who engaged in dialectical argumentation showed conceptual gains; those that engaged in mere one-sided argumentation or no argumentation did not gain at all.

In the present study, we attempted to extend these findings by further isolating the element thought to be responsible for the difference in cognitive gains, namely dialectical argumentation. Instead of being instructed to conduct an argumentative dialog with an assigned peer, students were prompted to engage in scripted monological dialectical argumentation on their own and a confederate’s solution of an evolutionary phenomenon in response to structural prompts read aloud by the confederate. In the control condition, subjects and the same confederate merely read aloud their solutions to each other, without discussing them further. Thus, students in both conditions were withheld from conducting a dialogical discussion with each other and were exposed to the same misconception in the intervention phase. The conditions were identical on factors such as, social facilitation, exposure to content matter and personification of viewpoints, and only differed in the elicitation of monological dialectical argumentation.

Coding Evolutionary understanding

The dependent measure of conceptual understanding concerning evolutionary change processes was assessed according to a coding scheme we developed in a previous study (Asterhan & Schwarz, 2005), which comprises two separate but complementary means of scoring:

**Explanatory schema score.** Based on and inspired by previous works we identified ten qualitatively different explanatory schemas1 in students’ explanations of evolutionary change. These different schemas were then quantitatively assessed on four different dimensions: Whether evolutionary change was considered at all, whether this change was explained, whether some sort of selection mechanism was used and whether existing intra-species variation was considered. Based on the appearance of each of these four dimensions, the ten qualitatively different explanatory schemas were assigned to one of five different categories. The score for each schema category was based on the number of dimensions that featured in the schema in that category. An additional null-category was added to distinguish between explanations that did not consider change (the lowest category) and those responses that simply did not answer the question at all (by stating that they did not know the answer or by repeating the data given in the item without providing additional information). This procedure yielded the following explanatory schema categories and their corresponding scores.

Non-answers (score 0): Responses that indicated absolute ignorance on the subject or responses that merely repeated the data that was given in the question item.

No change considered (score 1): Responses that simply denied that species evolve over time, responses that refer to a disaster scenario according to which everything will go extinct, and responses that refer to species intentionally moving away to other (better) areas to protect themselves.

Unexplained change (score: 2): This category includes responses that refer to species changing over time without providing a reason for it or describing how it occurs. Typically, they refer to evolution as a process according to which everything changes for the better and species keep on perfecting themselves.

Typological change: (score: 3) None of the four schemas in this category consider intra-species variation and selection, but they all acknowledge and explain change. Their accounts on why or how this change occurs is what distinguishes between them: Lamarckian (individual members acquire a trait which is then passed on to offspring), Mutations after (as a result of the change in the

---

1 Even though the term ‘mental models’ is the commonly used term in the concept learning literature, we prefer the term explanatory schemas (Ohlsson, 2002). While explanations are content-specific generative descriptions, explanatory schemas encode the structure shared by a set of explanations, that is defined by the set of generative relations and the way they are combined (Ohlsson, 2002).
environment, individual members “undergo mutations” which are then passed on to off-spring); Mating outside species (individual members will search to mate with members of other species that have an advantageous trait); and Dormant genes (the trait was always present in the species’ genetic make-up, but was dormant until it was “needed” and activated).

Hybrid explanations (score: 4) This category includes responses that integrate features of both natural selection as well as typological change mechanisms. They are similar to the former category in that they do not consider existing intra-species variation, but instead refer to variation that is created in reaction to environmental change and needs. However, they do consider selection: The transformed individual members of the species manage to survive and reproduce, whereas the others go extinct. Two different types of hybrid models were identified, differing in their accounts on how variation is created: (a) Some individual members of the species underwent genetic mutations in reaction to the change in the environment; and (b) Some individual members acquired the trait and this acquired trait is passed on.

Darwinian (score: 5) Responses that explain change in terms of natural selection and existing intra-species variation.

**Darwinist principles score.** The complementary coding scheme is adapted from Ohlsson (1992) who proposed a summary of modern evolution theory in the form of a number of key principles. It assesses the appearance of the following Darwinist principles: (1) Intra-species variability, Source of intra-species variability (i.e., random changes in genetic material), Differential survival rates, Differential reproduction rates, Accumulation of changes (i.e., the process is repeated many times), Changes within the population (i.e., proportion of individuals carrying the advantageous trait(s) will increase within the population).

The appearance of each principle in each target item was assessed according to the following grading key: The principle is mentioned and used correctly (2 points), only a part of the principle is mentioned and used correctly (1 point), or the principle is not mentioned at all (0 points).

Inter-rater agreement was \( r = .97 \) for the explanatory schema score and \( r = .95 \) for the Darwinian principles score.

**Coding argumentative reasoning**

Protocols of student conversations during the intervention phase were transcribed and coded according to the following criteria: Number of reasons participants proposed in favor of their own solution (strengths of their solution), number of objections they proposed against their solution (weaknesses in their solution), and number of supporting and objecting arguments for and against the confederate’s solution. Supports or objections that did not relate to content, but were of a pure evaluative nature without providing further reasons (e.g., “Your solution is better than mine”) or that related to superficial features of the solution only (e.g., “It is formulated very well”) were discarded. Any support or objection that did relate to content was coded, whether it was objectively correct or not. All twenty-two transcripts were coded by two independent human coders. Pearson correlations between their assessments were \( r = .81 \) (support own), \( r = .97 \) (objections own), \( r = .78 \) (support other) and \( r = .84 \) (objections other). Disagreements were resolved through discussion.

**Materials and procedure**

All students participated in the following sequence of activities: (1) Individual pretest (T1) to assess prior evolutionary understanding; (2) Instructional intervention: screening of instructional movie excerpt on Darwinist evolutionary theory; (3) Individual administration of a single transfer test item; (4) Experimental intervention with confederate in which they read each others' responses to the single test item according to different conditions; (5) Immediate posttest (T2): Re-administration of the single transfer item immediately following the experimental intervention (T2); (6) Delayed posttest (T3): Individual delayed post-test administered a week later. The total length of an experimental session was approximately one and a half hour.

All tests were administered to individuals as paper-and-pencil tests. The pre- and delayed posttest consisted of three open format test items (one warming-up item and two target items). Target test item templates were similar, but their content related to different evolutionary phenomena (i.e., running speed of cheetahs, mosquito insecticide immunity, webbed feet of ducks, pigmentation of moths and iguanas' swimming ability). Test item difficulty was assessed on a separate sample and was not found to differ among target items.

All students were shown a twenty-minute excerpt of an instructional movie that presented the Darwinian explanation of evolutionary processes. It mentioned a number of evolutionary changes in different Galapagos species, without providing any explanation of the process of change. Changes in a population of Galapagos finches, however, were discussed in detail and explained in terms of the Darwinian account, including a step-by-step graphical presentation.

Immediately after watching the movie excerpt, Ss were asked to solve a single test item on a novel evolutionary phenomenon (the webbed feet of Ducks) and hand in their answer sheets. They were then introduced to the confederate who, they were told, had just completed the same stages of the experiment as they had. The confederate was a female undergraduate of the same age group as participants. Great efforts were made to ensure that her physical appearance and behavior did not reveal her being a confederate.

Ss in the experimental conditions were told that they would participate in a short collaborative task with predefined roles (reader and respondent) and instructions. The confederate chose a piece of paper from an urn and
invariantly picked up the role of the 'reader'. All interactions were audio-taped.

The confederate then read aloud the general instructions that shortly described the task at hand and stated that the goal of the activity was to reach a better understanding of evolutionary change processes, rather than to 'win' a certain argument. The task itself consisted of nine steps that were read aloud by the collaborator. Six of these steps prompted the participant to engage in dialectical argumentative reasoning on their own and the confederate's explanation of the Ducks item; three of these steps contained procedural instructions:

First, participants were requested to read aloud their answer to the Ducks item from the previously filled out answer sheet. They were then asked to discuss the strengths of that solution, asked to critique it and discuss whether it explained the change that occurred to the Ducks' feet, in that sequence. Following, the confederate was requested to read aloud "her" solution, after which participants were asked to discuss that solution according to the previous steps. Finally, they were requested to evaluate the two solutions and their personal conceptual understanding of evolution. At each of the six argumentative steps the collaborator read the question aloud and waited for participants to complete their response, after which she then encouraged them to elaborate some more according to fixed prompts. At the end of the collaborative task, confederate and participant separately answered the Ducks item one more time on a clean answer sheet (T2).

The confederate's behavior was limited to reading the written prompts and reacting neutrally, but friendly, to the content of participants' responses. She was instructed to refrain from reacting to efforts from the participants to elicit comments or develop a dialogue.

So as to ensure uniformity concerning the exposure to another misconception (i.e., the confederate's solution), while, on the other hand, preserving a minimum difference between that and the S's explanatory schema, two different answer sheets were prepared for the confederate: One containing a typical Lamarckian answer and the other a typical explanatory schema of the same schema category, namely 'mutations after'. Both provided a typological explanation of evolutionary change, but differed in their description of the way the change occurred (see Coding section). While the dyad started reading the instructions the S's explanatory schema was quickly assessed based on his/her answer sheet filled out immediately following the movie. The confederate's response was then chosen according to the following principle: If the S's response represented a Lamarckian schema, the confederate was given the 'mutations after' response. In any other case, she was given the Lamarckian response. The experimenter then re-entered the room and "returned" the dyad's answer sheets in time for the S to read aloud his/her answer, with the excuse that the experimenter forgot to return them. The average length of scripted dialog was 8:05 minutes (ranging from 5:50 to 13:40).

In the control condition, Ss were also seated with the confederate. They were first given a short, neutral task (eight open questions on which animal is the fastest, strongest, heaviest, etceteras alive today). So as to control for time-on-task in both conditions, they were given seven minutes to complete this task. During this task the experimenter assessed the participant's mental model. Their sheets were then "returned" to them and they were requested to read their respective answers to each other while refraining from any further verbal communication, before answering the Ducks item once more on a new answer sheet (T2).

The confederate was given the Lamarckian response in 46% of the experimental interventions and in 50% of the control cases.

Results

Manipulation check
Nineteen of the twenty-two experimental subjects proposed at least one supporting reason for their own solution and an identical number of students proposed at least one objection. The number of subjects that mentioned at least one supporting argument for or at least one objecting argument against the confederate's solution was smaller (12 and 16, respectively). Two subjects only provided support for their own solution without further criticizing their own or the confederate's solution. The mean number of supporting arguments for own argument (M = 1.22, SD = .75), objections to own (M = 1.27, SD = .83), supporting arguments for other (M = .59, SD = .59) and objections against other (M = .91, SD = .87) were all significantly larger than zero (p < .001).

Conceptual gains
Mean explanatory schema scores and mean Darwinist principle scores were calculated for pretest, immediate posttest and delayed posttest occasions and are presented by condition in Figures 1 and 2.

Effect of condition Two separate, one-way analyses of variance (ANCOVA) were conducted on the delayed posttest variables of mean explanatory schema score and mean Darwinist principles score, while controlling for subjects' pretest scores on these respective variables and formal biology education as covariates. The experimental subjects showed superior conceptual understanding, both on the mean explanatory schema score, F (1, 38) = 8.87, p = .005, η² = .19, as well as on the Darwinist principles score, F (1, 38) = 9.81, p = .003, η² = .21. Previous biology education was not found to have an effect in either one of the analyses (F (1, 38) = 1.13, ns and F (1, 38) = 8.41, ns, respectively), whereas pretest scores were a significant predictor of delayed posttest performance in both (F (1, 38) = 35.30, p < .001 and F (1, 38) = 50.42, p < .001, respectively).
Patterns of change So as to uncover the patterns of change over the three tests in each of the conditions, the two dependent variables were subsequently analyzed in two separate analyses of variance (2X3) with time of assessment (pretest vs. immediate posttest vs. delayed posttest) as a within-subject factor, condition as a between-subject factor and formal schooling in Biology as a covariate. Huyn-Feldt corrections were applied when necessary. Previous biology education was not found to have an effect in either one of the analyses (F < 1 and F (2, 78) = 1.96, ns, respectively).

On the explanatory schema score a main effect for test occasion was found, F (1.70, 66.42) = 5.1, p = .012, η² = .116. More importantly, the effect of test occasion was found to be dependent on condition, F (1.70, 66.42) = 3.70, p = .037, η² = .087. Planned comparisons were conducted between the mean scores of each two consecutive tests (from T1 to T2 and from T2 to T3) within each condition: Cognitive gains of experimental Ss were obtained immediately following the intervention (F (1, 20) = 4.60, p = .044, η² = .19) and a similar level of performance was maintained on the delayed post test (F < 1). However, control subjects in this study showed a mere small and non-significant improvement at the immediate posttest (F (1, 18) = 1.89, ns), followed by a small and non-significant decrease at the delayed posttest (F (1, 18) = 2.00, ns). In other words, the explanatory schema score of control subjects did not significantly change from assessment to assessment.

With regards to number of Darwinist principles that were integrated in Ss' responses, time was not found to effect, F (2, 78) = 2.18, ns. However, condition was found to interact with time, F (1.70, 66.32) = 3.63, p = .039, η² = .09. The data in Figure 2 seem to suggest that the pattern of change over the three test occasions resembles the trends found on the explanatory schema score. However, none of the four planned comparisons reached significance.

Conceptual change

Our hypothesis that the engagement in monological argumentation improves conceptual understanding was confirmed at the explanatory schema level as well as the number of Darwinist principles. The question remains, however, whether this advantage will also hold under the stricter criterion of conceptual change. Whereas the mean increase for the argumentative condition was .36, the intra-group variance was relatively large: Some radically changed their explanatory schemas, whereas others did not gain at all, or even slightly deteriorated. How many students did actually attain conceptual change? To answer such questions we first have to define what is considered conceptual change in evolutionary theory and what would account for a 'mere' amelioration of existing conceptions?

Whereas the surface features of two explanatory schemas from the same category are different, the underlying explanatory mechanism is the same. An intra-categorical shift from one schema to another would, therefore, not account for conceptual change. We therefore argue that only shifts from one explanatory schema category to another involve the substantive re-organization that is described in conceptual change theory and research (e.g., Vosniadou, 1999).

However, the student responses to the two target items on both the pre- and the posttests reveal that students were not necessarily consistent in the explanatory schemas they applied on a given test occasion: Only seventy percent of the total number of pretest and posttest questionnaires revealed consistent use of explanatory schema categories on a given test (76% and 64% on the pre- and delayed posttest, respectively). The mean difference between two test item responses on 'inconsistent' questionnaires was similar on pre- and delayed post-test (M = 1.00, SD = .58 and M = 1.01, SD = .49, respectively).

A definition of conceptual change has to take this instability into account. Whereas a student who applied an explanatory schema of a one-point higher category on only one of the two test items (i.e., a mean pre- to posttest increase of .5 points) has indeed shown improved conceptual understanding, we argue that this does not provide sufficient proof for a substantive change. We therefore defined an increase of at least one point from mean pretest to mean posttest score as sufficient proof for a more profound change in conceptual understanding, that is: conceptual change. Table 1 presents the number of students...
for which sufficient evidence for conceptual change from pre- to delayed posttest was found. Students with pretest scores above 4.0 were excluded from this analysis.

Table 1: Number of students for which evidence of conceptual change was found, by condition.

<table>
<thead>
<tr>
<th></th>
<th>No conceptual change (d &lt; 1)</th>
<th>Conceptual change (d &gt;= 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>Control</td>
<td>15</td>
<td>0</td>
</tr>
</tbody>
</table>

* d = difference between mean pre- and delayed post-test scores

The data in Table 1 show that among those students that could achieve conceptual change (i.e., pre-test scores below or equal to 4.0), forty percent of experimental students and none of the control subjects substantively changed the explanatory schemas they used to explain evolutionary phenomena. The occurrence of conceptual change was found to be dependent on experimental condition, Fisher's exact test, \( \chi^2 (1, N = 30) = 7.50, p = .017 \).

Discussion

The results presented here provide further support for the assertion that argumentation promotes conceptual understanding in a cognitive conflict learning paradigm. Students that were engaged in scripted dialectical argumentation by providing justifications, counterarguments and evaluations on their own and another person's solution, showed superior conceptual understanding in evolution and were more likely to have attained conceptual change, than control students. This study also replicated the pattern of change in argumentative condition which was characterized by immediate conceptual gains that were preserved until the delayed posttest (Asterhan & Schwarz, 2005). Students in the control condition, however, did not show improvement on any of the test occasions. When they were refrained from engagement in dialog and were exposed to the same misconception, the pattern of temporary gains found in the previous study disappeared.

The finding that at least a third of students in our sample applied substantively different explanatory schemas at the same test occasion is especially noteworthy. It, first of all, stresses the importance of multiple assessments. Secondly, this finding seems to corroborate with current theoretical models and empirical findings that regard explanatory schemas (mental models) as temporary mental representations that are constructed in working memory according to specific demands of the task at hand and based on schema configurations stored in LTM (Schnotz & Preuss, 1999; Ohlsson, 2002). At any given time, different schema configurations compete with each other and have different likelihoods of being activated. As a result, a student may construct different mental models to explain different phenomena that relate to the same scientific concept. Accordingly, conceptual change should not be conceptualized as a shift from one particular conception to the other. Like Siegler's overlapping waves theory (1996), it should be regarded as a change in the probability according to which more advanced schema configurations are activated and used to construct explanations.

Acknowledgments

This work was partially funded by a grant from the Israel Foundations Trustee (No. 27/35) to Christa Asterhan.

References


